

"Get It, Catalog It, Promote It": New Challenges to Providing Access to Special Collections

Beth M. Whittaker

It is becoming a cliché to state that as research library collections become more homogenous—through the consolidation of purchasing (such as widespread, availability of the same electronic serials) and through the increase in digital content accessible to library users regardless of institutional affiliation—special collections will be what sets apart one library from another. This prediction has become so common that it even figures in published library satires.¹ Administrators seem to be hearing this message as well. In his plenary address at the 2005 RBMS Preconference, Steven E. Smith, Director of the Cushing Memorial Library and Archives at Texas A&M University, urged special collections professionals to take advantage of this potential golden age.² If this is the case, and I sincerely hope it is, we must also focus attention on the technical services that make these collections usable.

"Access" to special collections material has traditionally depended on those technical services functions that occur "behind the scenes": acquisitions, cataloging, and preservation. If material is too fragile to be consulted, or if the only person who knows of its existence is the curator who happens to be out of town, patrons cannot "access" it in any sense of the word. As the use of off-site storage for all library materials increases, technical service functions are becoming more necessary for "regular" collections as well, since these stored materials can no longer be visually browsed nor can reference librarians easily access them.

It is very difficult for anyone working in special collections to argue that there is not an "access problem" in our field. Not only do we see physical backlogs of completely uncataloged material and "underlogs" of material with inadequate access, but we also are aware that in many cases there are limitations to what we can do to provide intellectual access to material. Lack of experience with specific languages, scripts, genres, or subject matters can frustrate staff working in many institutions, including those that provide traditional item level full MARC cataloging.

We have to use some word to describe the physical and virtual tools we develop to connect materials and users, and although the term "access"—even when applied strictly to a library context—already has multiple meanings, it is probably the best word at hand. I use the equally unwieldy term "modes of access" to refer to all those tools that provide intellectual access to our collections by pointing people to the resources they need (such as catalog records, finding aids, card files, and databases). But we should remember that these modes of access may not necessarily be used by our patrons, since we have a long tradition of creating in-house tools that serve the same purpose. Because many of our initiatives to improve access have been aimed at translating these mediated tools to a publicly accessible format, much of the information can still remain "behind the scenes." Collections databases, for example, are often built with public and staff views, each of which display varying levels of detail.

It is worth highlighting a few comments made by Daniel Traister, Curator of Research Services at the University of Pennsylvania's Rare Book and Manuscript Library, about expanding access to special collections. Traister argues that special collections have undergone a "major shift in ... attitudes" and are "more welcoming" to users. The historic attitude of "get it, catalog it, preserve it" (the classic technical services functions) has become "get it, catalog it, promote it." He acknowledges a climate of economic scarcity, and underscores the emerging importance of classrooms, the Web, and seminars in special collections work.³

What should this mean for technical services? We must ask ourselves if our modes of access are really welcoming and user-friendly. And if so, are they friendly for all types of users, or just those bibliographic experts with whom we have traditionally been comfortable working and for whom we have designed much of our infrastructure? How can our modes of access be used to promote our collections? How can we "justify" our work, and should we volunteer to do so or wait until we are forced to do so by fiscally conservative administrators? Finally, what role do technical services play in teaching and outreach, even if catalogers and other technical service staff are not the ones who do the actual teaching or reach out from our isolated cubicles?

Although my experience has been mostly with large academic library special collections, I suspect many of the same factors influence other environments, too, and that "technical services" duties are generally similar in nature if not in specifics.

The same does not necessarily hold true for the process of acquiring materials for special collections, however. Except for the significant changes brought about by the growth of Internet buying and selling, little can be said in a general way about the acquisition process, since it is highly dependent on genre, collection, institution, and vendor. Cataloging and preservation, which I will address much more thoroughly, take place after the material is already acquired. I will also address to some extent digitization efforts that are not directly related to preservation, since these, too, can be considered as a part of technical services.

In an attempt to establish my cataloging *bona fides*, I must admit that I have in the past been guilty of designing catalog records to some extent for their own sake. Constructing a perfect MARC record, complete with rarely used indicators and every possible note, used to thrill me. Nowadays, perhaps as I have taken on additional duties, I have much less tolerance for the time and effort involved in taking a description from "good enough" to "perfect." I am also coming to believe that we can never achieve perfection, even with all the time and expertise in the world. Any mode of access (MARC record, finding aid, etc.) may be a thing of beauty and elegance, providing every conceivable piece of information and meshing perfectly with existing records in the catalog, but its only actual purpose is to link users and materials. Since we can never really anticipate what users will be looking for, the extra effort required to create a "perfect" record is (to my mind) often unjustifiable.

Thinking Globally

By now, the "Exposing Hidden Collections" initiative of the Association of Research Libraries' Special Collections Task Force is likely a familiar one to most in the field.⁴ This work underscores the depth and breadth of "the access problem." High-profile initiatives often get the attention of library administrators, and this is a good thing. I agree our backlogs are unsustainable and, often, unjustifiable. But a larger issue is how to provide a "strategy of access" to all special collections in a given environment, an issue that no amount of grant-funded or cooperative cataloging projects can solve in a vacuum.

We have created (sometimes with deliberate intent) a double or triple standard: large backlogs of undocumented material coexist with detailed information about other parts of the collection, often as item-level access to digital images. This decision can be rational and in the best interests of the collection, but everyone (curators, public services, and technical services) needs to understand the priorities and the rationale behind that decision.

In the 1990s, the phrase "technical services" seemed to indicate that catalogers should be the library staff handling sexy new "metadata." It was "technical," after all. And in special collections, the future was even "sexier" because we were often early adopters of digital initiatives. Also, special collections that had often been part of larger library systems and forced to compromise their local practices through shared catalogs (and often disastrous retrospective conversion) finally began to realize some degree of independence. With a few savvy volunteers or student assistants, we could create our own Web pages, link to HTML finding aids, mount our own collections databases, and essentially present a very public face for our special materials and services.

I am ashamed to say that, in many cases, technical services librarians have not lived up to the promise of these new technologies.⁵ From the start, we could have integrated "metadata" functions and digital initiatives into our existing workflows, but the landscape at the beginning of the 21st century still seems to be one of fragmented responsibilities. Professionals working with these promising new initiatives are often not part of the traditional technical services work of special collections; rare book catalogers have *not* become metadata Librarians. For autonomous special collections libraries and archives, the situation may be brighter, but in large academic institutions, I see evidence of increasing fragmentation among a number of communities as technology progresses.

Largely because we have been unable to absorb all these emerging functions, providing access to the digital world now calls for more cooperation with information technology staff and others outside special collections. In my experience, special collections librarians cannot agree among ourselves about how to best enhance access. Consider starting a discussion with colleagues on the following points: What is the usefulness of bibliographic collations in rare book records? What about the proper number and function of tracings for contributors or minor "authors?" How helpful are special homegrown classifications schemes? Even within our own communities, special collections librarians do not have a strong record for cooperating.

The ongoing efforts of our cooperative cataloging projects to use the Web to expose our records, which were created to function in a given institutional context, highlight the potential weaknesses of our collaborations. OCLC's Open WorldCat is receiving press for releasing MARC records from OCLC's catalog to search engines.⁶ Users can now discover these records through the usual Web-searching routines, and eventually, perhaps, locate individual copies in libraries (or, of course, from booksellers).

The Open WorldCat project has been trumpeted in library circles as a way to promote our collections to users unwilling or unable to search a separate catalog. But for special collections records, it is perhaps more problematic than it seems. First of all, catalogers working in special collections have consistently kept in mind that we do not only create records; we create catalogs. Decisions about access points, notes, and other features are made in the context of a given catalog with a particular user base in mind. Even if we have not always correctly identified the user base, at least we have attempted to do so. Taking these records out of context could lead to unintended confusion.

Even more troubling is the fact that the entire OCLC structure is based on one master record, which negates the level of detail we have often put into records in our local catalogs.

Donor notes, copy-specific details, and preservation information are clearly unnecessary in a "Google" view of a record, when users want only the information contained in a given text. As administrators begin to wonder whether we even need an institutional catalog once all the records are on Google, we must keep this in mind.

A similar initiative is RedLightGreen, an attempt to entice researchers to locate records from the Research Libraries Group (RLG) Union Catalog through an easy-to-use search interface.⁷ These are exciting initiatives for the library world, but we must remember that one size does not fit all in the MARC world, and even the advanced "more search options" search box offered by RedLightGreen does not accommodate the kinds of searches we have often expected of our users in special collections and that have, therefore, informed how we catalog materials.

Of course, the efforts extend far beyond simply creating MARC records, which have traditionally been used to describe books. There are also finding aids or inventories and other types of records that are used for archives, manuscript collections, and objects. We now face a situation where even in a small library, we have information about our collections existing in catalogs, databases, documents (both print and electronic), legacy card files, and many other types of systems. Sometimes it is an effort even to remember where to look, let alone how to search or what the idiosyncratic features of a particular system might be.

Have we fallen in love with the illusion of increased access, simply because we have multiplied the places where information about our collections lives, and the technological methods used to display this information? Is this the source of the access problem? If so, simply "retroconning" these systems to make them publicly searchable will not help us at all.

How are access problems today different from those we faced in the past? If your library is like all the ones where I have worked, you used to have a pretty good idea where "the backlog" was: in a physical location, often the basement, or otherwise out of sight. We also knew which collections had finding aids because that information was kept in binders lined up on a shelf. I doubt we were even happy with these limitations, but we knew the scope of the problem and with luck and a good memory, most staff would be able to assist users in their research. The prospect of a grant-funded project or a talented volunteer always meant that organization and access lay tantalizingly just around the corner.

Today, of course, we no longer control all the haystacks in which to go searching for needles. People do not have to come to the reading room to see our idiosyncratic descriptions and finding aids, even for research purposes, let alone for copying and using images. They can, to some extent even with special collections, have unmediated access, particularly when their needs do not infringe on our cost recovery efforts.

Perhaps more disturbing, we are not really sure any more what we should consider to be a "backlog." For example, I serve as editor of the controlled vocabularies maintained by RBMS's Bibliographic Standards Committee.⁸ Guided by requests from professionals working with collections, we are constantly expanding these thesauri to accommodate changing patterns of scholarship and the need for new descriptive terms. Again, this is a good thing. But once catalogers have added these controlled terms to records, thereby allowing searching in new and useful ways, we still have work to do.

Either we must explain to our colleagues (curators and public services staff) and through them, to our patrons, why not all genres (or binding terms or what have you) are indexed for all materials in all our collections to the same level, or we must try to implement policies about what to trace and when, and then do significant retrospective work to implement these policies consistently. A third option, and one I suspect is more common, is simply to give up and console

ourselves that at least researchers can find this type of thing from here forward ... or at least until the controlled vocabulary terms change or until the next wave of scholarship overtakes our collections. Even within a relatively small community, we cannot meet all the expectations of our users (even when these users work in our own institutions), nor can we be assured that the left hand and the right hand know what the other is doing.

This is not just the case for traditional descriptive information. What about the information librarians create—such as exhibition catalogs, articles, and bibliographies? How do we provide access to these "hidden collections?" What about the information we are creating in the process of doing our technical services work? In a discussion among special collections catalogers about how to handle material found in books, it was noted that by removing laid-in materials and placing them in some sort of "provenance file" without creating adequate access to the materials and sufficient links between the "stuff" and the books, we were in fact creating brand-new "hidden collections" even as we cataloged individual items with MARC records.⁹ It is enough to make one despair of ever solving the access problem. Just as curators are no longer the "gatekeepers" once the digital material is "out there," catalogers are no longer "gatekeepers" of description or intellectual access. In truth, catalogers may be even less comfortable with this arrangement.

Acting Locally

By way of example, I now present a few ideas of practical approaches at the library level that may be instructive. The first strategy of enhancing access is to build on what we have. Due to retrospective conversion projects and a robust shared cataloging tradition, most institutions have pretty solid information in our integrated library systems for a lot (if not most) of our print collections. We can use some of the functionality of these systems to provide access to special collections in new ways, or to mirror traditional means of enhanced access more efficiently and inexpensively.

A simple example, depending on how one's system is set up, is to use a MARC field for collection names. (In many systems, the best way to do this would be to add a "title" using a local note, such as a 793 local added-entry uniform-title field.) In a Web interface, this title becomes a hyperlink to a dynamically updated list of material in a given collection, meaning that curators no longer need to maintain separate files—the need for paper collection files is only as great as one's skepticism about integrated library systems. It can also serve as a highly visible way to recognize donors.

This type of creative manipulation of MARC can also be used to replicate many of the special files long maintained by libraries for researchers. One of my first jobs in a library was to type provenance tracings at the top of OCLC-produced cards for filing in an actual card catalog. While some libraries no doubt still rely on this approach, the increasingly complex needs of the larger library community have ensured that at least some of our "special" needs are met by vendors. It is encouraging that of the recommendations put forward in 1998 by the Bibliographic Standards Committee of the RBMS for improving library systems, most have now been met by the larger vendors.¹⁰ Of course, cost is always a factor, and while student assistants tend to earn very low wages, we must acknowledge that even low-tech efforts to improve access come at a price.

Digitizing Our Collections

Although the initial excitement over the paperless library seems (fortunately) to have died down in recent years, we still often hear talk of providing access to special collections by just digitizing the darned stuff. Certainly digitization of print, in particular, is easy enough; the technology is constantly improving so that our most precious volumes can be digitized more safely, more quickly, and more inexpensively than ever. Publicity material for one such process affirms that "large-scale digitization of even the most fragile books is now both easy and affordable while maintaining the integrity and safety of the documents."¹¹

But digital images of our collections are not the end product, and one anecdote should serve to illustrate the scope of the problem for those who have been so fortunate as to not have encountered it. At the Ohio State University's Rare Books and Manuscripts Library, we were able to take advantage of preservation treatment to scan a 1563 edition of John Foxe's *Actes and Monuments*—known popularly as *The Book of Martyrs*—part of a prominent collection of this landmark work. The book had been disbound for cleaning and repair and is a perfect example of the relationship between preservation and access. This edition is over 1,700 pages, and is 33 cm. tall. Digital photography was undertaken at a very high resolution, but "use copy" JPG files were also derived from the original photographs, and in this format, all the pages fit on one CD.

At one level, just creating this CD of lower quality images provides more access. Patrons can currently check out the CD from the library and take it home to view in any computer with a CD drive. Additional copies could be easily made available, meaning that even when the reading room is closed, or when the physical item is being used in the reading room, multiple patrons could view "pretty good" facsimiles of the text. This is, in fact, increasing access to our special collections.

However, we have not been able to deliver the "digital preservation" TIFF images in a similar way due to the size of the files; we cannot deliver them across the Web in any meaningful way. Consider whether one can realistically view over 1,700 pages of high resolution scans or photos at one's desktop. Can scholars search for certain words or phrases in this work, one TIFF image at a time? Can art historians compare the woodcuts of martyrs being burnt at the stake in this text with those in subsequent editions of the work? Some might argue that it would be easier and quicker to flip through the pages of the physical volume.

While sophisticated and easy-to-manipulate page images are better (in some ways) than working in a reading room, they are not useful enough. For researchers truly to benefit from the technology, we must somehow break up the contents into usable and significant bits. What are these bits—chapters, pages, paragraphs, words? Might not the definition of "usable bits" change with scholarship, the shifting focus of our collections, and changes in technology? The last frontier for "technical services" may be searchable full-text. Now, we can provide this even for difficult-to-read fonts and fragile pages, either through OCR of text or rekeying and markup by library staff or vendors. Technology is not the problem, but the allocation of resources is. Communication among all those involved in such an endeavor is often a very big problem as well.

Margaret Mitchell, Associate Professor of New Testament Studies at the University of Chicago Divinity School, has spoken to the RBMS about the creation of a flexible and complex digital presentation for a highly studied manuscript of St. Mark's Gospel.¹² The "ooh and ah" factor for this kind of work is high, but can such an amount of work realistically be applied to more than a few high-interest volumes in most collections?

Perhaps you have similar projects at your institution, living in an access limbo awaiting more time, more money, or better technology. Meanwhile, our administrators are reminding us that somehow Google is, in some magical proprietary process, (apparently) being successful at what we struggle with, and those administrators consequently question our cost estimates and the amount of work involved.

Collection-Level Cataloging

Just as improvements in digitization technology have yet to lead to the elimination of paper volumes in libraries, the promise of collection-level cataloging has not led to the elimination of our backlogs. Nervous catalogers, convinced that collection-level cataloging was a ploy to send them to the unemployment line, can relax for the time being. Yet I believe there is a prominent place for collection-level records in the access strategy of special collections institutions.

Again, an example from my own institution may be informative. The San Francisco Academy of Comic Art collection was compiled by collector Bill Blackbeard over several decades. In his desire to collect "popular narrative art," Blackbeard eventually filled fourteen rooms of a two-story, 4,000-square-foot house with formats that include newspapers and periodicals, dime novels, comic books, fanzines, and reference works. When the collection arrived in Columbus, it filled six semitrailers and weighed in at seventy-five tons. We should acknowledge that there is no collection-level record in the world that can adequately describe seventy-five tons of material.

Instead, the project archivist and I developed a strategy to provide access to several parts of the collection in different ways. Monographs and serials were cataloged using item-level MARC records, for most of which, luckily, copy cataloging was available. Added titles (using the MARC 793 local field) collocated all the records that formed part of the collection, making additional record-keeping unnecessary. The business records for the organization were described using a brief record for all ninety-two boxes. While an inventory of these boxes is available in the Cartoon Research Library, it is not yet available to the public.

Finally, the archivist spent the most time on the newspaper sections, a particular strength of the collection and an area not widely collected. A MARC record for the subcollection links to the finding aid. This is no average inventory, but a complex finding aid developed with the specific nature of the collection in mind. It includes title-level access to individual comics features, like *Faithful Jems*, *Barney Google and Snuffy Smith*, or *Peanuts*; or to newspaper names represented by comics sections, such as the *St. Louis Republic* or the *Nashville Banner*. The list of comics features links to PDFs that provide chronological listings of holdings, while additional indexes are provided for individual artists and comics sections.¹³ The use of this process meant that students could work on a simple spreadsheet interface while the archivist continued to work on an encoded finding aid.

For the archivist, this procedure was not only the best way but the only way to go. It also made sense from a traditional library cataloging approach. Monographs and serials might very well be sought through known item searches (author, title) or by subject or genre. Cataloging these items, although time-consuming, was possible and worthwhile. Newspaper sections might be looked for by comics scholars, but the fact that individual newspapers are listed means that scholars doing other types of work can find relevant material as well. Unfortunately, there are parts of this collection that still have not been made "accessible." We have, in good library tradition, left something for our successors to do.

Preservation

Finally, I would like to address preservation explicitly. I am particularly interested in the false dichotomy between "preserving the stuff" and "preserving modes of access to the stuff." (Although not particularly descriptive, "stuff" captures the vague and fascinating nature of our work.) Are digital images within a database considered "stuff," or "modes of access"? I would argue both.

Preserving the "stuff" falls somewhere in a long tradition of surrogates and facsimiles, designed to accommodate a perceived need for wider distribution than the original works would allow. Preserving "modes of access" is, however, often overlooked. CDs can be stored off-site for security, but of course one must ensure access to the file formats and perhaps also the software used to create the data. Card files and databases can be reformatted. And what about troubleshooting access—how does one ensure that links to a finding aid keep working, for example?

Should we be concerned about this, or is this someone else's problem? I believe the attitude that builds a wall between preservation and access, between librarians and information technology staff, highlights the fragmentation I discussed earlier. Technological solutions are only means, not ends. Technology does not make decisions about how, when, and why to implement itself, and while IT professionals are valuable allies, we are selling our collections and our professional expertise short if we leave all these decisions to someone else.

Conclusion

One could argue that by creating these different modes of access, we have only perpetuated the problem of multiple card files and databases that I mentioned previously, and that would be right. But I believe that by using, as much as possible, the power of our ILS and our MARC traditions, we have at least made sure there is a home for most types of access for the short term. Another example of the autonomy brought by digitization is that special collections can control, to a large extent, what is linked out, even if the records that contain these links must conform to a larger set of guidelines. In fact, in some institutions, special collections cataloging workflows have served as models for other nonstandard use of MARC linking, and Encoded Archival Description (EAD) may be a library's first introduction to non-MARC metadata. It is always gratifying when a lesson learned in special collections turns out to have broader impact.

At the same time, creating links between different types of descriptions serves, at least in a technical way, to strengthen the associations between the different descriptive traditions embodied in cultural heritage collections. No longer must we make decisions about following library cataloging *or* archival processing; we can provide both types of descriptions with little or no additional effort. Given the ability to map many different types of metadata to intermediate XML schemes or to allow "harvesting" of metadata across collections, this approach further integrates our collections and hopefully makes it easier for users to discover our resources, while maintaining the distinctions that have developed, with good reasons, among different traditions represented in our collections.

It has been also argued that MARC itself is useless for the hyperlinked and varied information world we live in, and that we should be migrating our work to other schemes, "throwing the dirt on the coffin of MARC" just as soon as we can—a thought that segues nicely into a few cautionary words for those of us insulated and surrounded by special collections.

It does us no good to isolate special collections from larger access initiatives in the library community, insisting our "stuff" is so darned special we can ignore the rest of the library world.¹⁴ While we may be frustrated with the glacial pace of revisions to the Anglo-American Cataloging Rules, for example, it is worth our while to keep our eyes on the development of RDA, *Resource Description and Access*.¹⁵ We must also acknowledge that the much smaller special collections community has not historically shown itself to be particularly flexible or nimble either.

Instead, we can be test cases for new initiatives. Those of us working in large academic libraries have no doubt heard more than we care to about institutional repositories (IR). We should also be striking while this particular iron is hot. IR systems were designed for preprints and other grey literature, or as storehouses of research data, not for special collections. Nonetheless, at Ohio State archivists are spearheading a project to contribute oral history transcripts to our repository.¹⁶ We are able to contribute to the development of this open-source software so that ultimately, it will be more useful for these types of collections and will provide increased access to our archival materials.

Recently, I have become more and more concerned about the fragmentation of the cataloging world and the future of technical services both within special collections and libraries in general. While I realize I have made many generalizations about our work and our professional environments, perhaps this call to arms will challenge readers to prove me wrong, through an increased commitment on the part of all those who work with special collections, to address our access problems in a systematic and collaborative way, both in our own institutions and the international library world.

Notes

1. Norman D. Stevens, "The Fully Electronic Library," *College & Research Libraries* 67:1 (2006):9.
2. Steven Escar Smith, "From 'Treasure Room' to 'School Room': Special Collections and Education," *RBM* 7:1 (2006): 31-39.
3. Daniel Traister, "Is There a Future for Special Collections? And Should There Be? A Polemical Essay," *RBM* 1:1 (2000): 54-76.
4. See particularly the Task Force's homepage (<http://www.arl.org/collect/spcoll/>) and the special issue of *RBM*, 5:2 (2004) devoted to the initiative.
5. For this argument, see Christine DeZelar-Tiedman, "Crashing the Party: Catalogers as Digital Librarians," *OCLC Systems & Services* 20:4 (2004): 145-47.
6. Project Web site: "www.oclc.org/worldcat/open."
7. RedLightGreen homepage: www.redlightgreen.com.
8. "RBMS Controlled Vocabularies for Use in Rare Book and Special Collections Cataloging," http://rbrns.irifo/conmuttees/bibliographic_standards/controlledvocabularies.
9. Discussion at MARC for Special Collections Discussion Group, American Library Association, 2004.
10. "Guide to Rare Book Records in Online Systems," www.folger.edu/bsc/guide.html.
11. Kims Technologies: www.kirtas-tech.com/.
12. Archaic Mark MS 972: <http://goodspeed.lib.uchicago.edu/archaicmark/home.html>.
13. "Guide to the San Francisco Academy of Comic Art Collection: Newspaper Comic Strips" finding aid: http://cartoons.osu.edu/finding_aids/sfaca.
14. For examples of such initiatives, see *Functional Requirements for Bibliographic Records* (www.ifla.org/VII/s13/frbr/frbr.htm) and the "Open Archives Initiative Protocol for Metadata Harvesting" (www.openarchives.org/OAI/openarchivesprotocol.html).
15. "RDA: Resource Description and Access" page: www.collectionscanada.ca/jsc/rda.html.
16. See "Polar Oral History Community" at: <https://kb.osu.edu/dspace/handle/1811/6039> and "OSU Oral History Community" at: <https://kb.osu.edu/dspace/handle/1811/403>.